ISR3
71

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 July 2001 (26.07.2001)

PCT

(10) International Publication Number
WO 01/53539 A1

(51) International Patent Classification[7]: C12Q 1/68, C12P 19/34, C07H 21/00, 21/02, 21/04

(21) International Application Number: PCT/US01/00291

(22) International Filing Date: 4 January 2001 (04.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/177,873 24 January 2000 (24.01.2000) US

(71) Applicant: **PHYLOS, INC.** [US/US]; 128 Spring Street, Lexington, MA 02421 (US).

(72) Inventor: **WAGNER, Richard**; 1007 Lowell Road, Concord, MA 01742 (US).

(74) Agent: **ELBING, Karen, L.**; Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110-2214 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/53539 A1

(54) Title: SENSITIVE, MULTIPLEXED DIAGNOSTIC ASSAYS FOR PROTEIN ANALYSIS

(57) Abstract: Disclosed herein are methods for detecting multiple compounds in a sample, involving: (a) contacting the sample with a mixture of binding reagents, the binding reagents being nucleic acid-protein fusions, each having (i) a protein portion which is known to specifically bind to one of the compounds and (ii) a nucleic acid portion which encodes the protein portion and which includes a unique identification tag; (b) allowing the protein portions of the binding reagents and the compounds to form complexes; (c) capturing the binding reagent-compound complexes; (d) amplifying the nucleic acid portions of the complexed binding reagents; and (e) detecting the unique identification tag of each of the amplified nucleic acids, thereby detecting the corresponding compounds in the sample. Also disclosed herein are kits for carrying out such methods.

SENSITIVE, MULTIPLEXED DIAGNOSTIC ASSAYS FOR PROTEIN ANALYSIS

Background of the Invention

In general, the invention relates to diagnostic methods involving multiplex analysis.

A variety of methods exist to detect multiple species in a biological sample. These include ELISA based immunoabsorbent assays, protein biochips, and the like. Each of these methods suffers from limitations in detection sensitivity or selectivity, due, for example, to kinetics of binding or sensitivity of detection reagents. In addition, these techniques are also limited in terms of the number of molecules that can be rapidly detected.

Summary of the Invention

The present invention involves a novel multiplex diagnostic approach for the ultra-sensitive detection of molecules in biological samples.

According to this approach, one begins with a set of uniquely defined high affinity binding reagents (typically protein binding reagents). Each of these reagents binds to a different target in a sample, facilitating the detection of several targets simultaneously. The targets of the binding reagents are frequently proteins, but they may be any moiety capable of specific binding, including, for example, nucleic acids or sugar moieties. Such binding reagents may represent naturally-occurring or partially or completely synthetic amino acid sequences. Examples of naturally-occurring binding reagents include, without limitation, members of the following binding pairs: antigen/antibody pairs, protein/inhibitor pairs, receptor/ligand pairs (for

example cell surface receptor/ligand pairs, such as hormone receptor/peptide hormone pairs), enzyme/substrate pairs (for example, kinase/substrate pairs), lectin/carbohydrate pairs, oligomeric or heterooligomeric protein aggregates, DNA binding protein/DNA binding site pairs, RNA/protein pairs, and nucleic acid duplexes, heteroduplexes, or ligated strands, as well as any molecule which is capable of forming one or more covalent or non-covalent bonds (for example, disulfide bonds) with any portion of a nucleic acid-protein fusion. In addition to naturally-occurring binding partner members, binding reagents may be derived by any technique, for example, by directed evolution approaches using a desired protein as the binding target.

Whether naturally-occurring or synthetic, when mixtures of binding reagents are utilized in a single diagnostic reaction mixture, they are preferably similar in composition and amino acid length. In a particularly preferred approach, one starts with a common amino acid scaffold or structural motif that displays the binding domain on one face of the molecule, as is the case for an antibody scaffold that displays CDR regions as a binding region of the molecule. A particularly useful binding scaffold is the 10^{th} domain of type III fibronectin (see, for example, Lipovsek et al., Protein Scaffolds for Antibody Mimics and Other Binding Proteins, U.S.S.N. 09/456,693; U.S.S.N. 09/515,260; U.S.S.N. 09/688,566; WO 00/34784).

The compound to be detected by the present approach may be any substance to which a protein may bind, and is preferably itself a protein. Such target compounds may be present in any sample, for example, any biological sample. Typical biological samples include, without limitation, any fluid or tissue derived from an organism, for example, a plant or a mammal such as a human.

As a central feature of the invention, each of the binding domains is covalently attached to a nucleic acid that encodes the binding domain. Such nucleic acid-protein fusion molecules can be produced by any method, for example, by the method of Roberts and Szostak (Szostak et al., U.S.S.N. 09/007,005 and U.S.S.N. 09/247,190; Szostak et al., WO 98/31700; Roberts & Szostak, Proc. Natl. Acad. Sci. USA (1997) vol. 94, p. 12297-12302) using a peptide acceptor, such as puromycin, as a covalent linking agent. As used herein, by a "peptide acceptor" is meant any molecule capable of being added to the C-terminus of a growing protein chain by the catalytic activity of the ribosomal peptidyl transferase function. Typically, such molecules contain (i) a nucleotide or nucleotide-like moiety (for example, adenosine or an adenosine analog (di-methylation at the N-6 amino position is acceptable)), (ii) an amino acid or amino acid-like moiety (for example, any of the 20 D- or L-amino acids or any amino acid analog thereof (for example, O-methyl tyrosine or any of the analogs described by Ellman et al., Meth. Enzymol. 202:301, 1991), and (iii) a linkage between the two (for example, an ester, amide, or ketone linkage at the 3' position or, less preferably, the 2' position); preferably, this linkage does not significantly perturb the pucker of the ring from the natural ribonucleotide conformation. Peptide acceptors may also possess a nucleophile, which may be, without limitation, an amino group, a hydroxyl group, or a sulfhydryl group. In addition, peptide acceptors may be composed of nucleotide mimetics, amino acid mimetics, or mimetics of the combined nucleotide-amino acid structure. As noted above, puromycin represents a preferred peptide acceptor for use in the present method.

In addition to covalently bonded RNA-protein fusions, any other unique, PCR-amplifiable nucleic acid (for example, RNA, DNA, PNA, or any other nucleic acid which includes two or more covalently bonded, naturally-

occurring or modified ribonucleotides or deoxyribonucleotides) can be coupled covalently or non-covalently to each individual binding domain. The protein portions of the fusions are typically composed of naturally-occurring amino acid residues, but may also include amino acid analogs or derivatives, joined by peptide or peptoid bond(s).

Of particular importance is that each binding domain is associated with (and can therefore be identified by) a unique, amplifiable nucleic acid tag, and that each tag in a multiplex reaction is of identical (or essentially identical) length to avoid amplification (for example, PCR) biases. Such unique identification tags are nucleic acid sequences that differ sufficiently in sequence from other tags in a given population or reaction mixture that significant cross-hybridization does not occur under the conditions employed. These unique identification tags may be present in the protein encoding portion of the fusion (for example, the tag can be a randomized portion of the protein scaffold, such as a randomized loop of the 10^{th} domain of fibronectin type III). Alternatively, the unique identification tag can be added to the nucleic acid portion of the fusion molecule and be positioned outside of the nucleic acid sequence which encodes the compound binding domain or, if present, its associated scaffold region. In the latter case, unique identification tags may be chosen which most effectively, most selectively, or most conveniently identify the fusion molecule. For example, if binding reagents are deconvoluted on a DNA chip, tag(s) may be chosen which best hybridize to immobilized chip nucleic acid(s) or which are compatible with commercially available chip arrays. Although DNA chips represent a preferred solid support according to the invention, deconvolution may also be carried out on other solid substrates including, without limitation, any other type of chip (for example, silica-based, glass, or gold chip), glass slide, membrane, bead, solid particle (for example,

agarose, sepharose, or magnetic bead), column (or column material), membrane (for example, the membrane of a liposome or vesicle), test tube, or microtiter dish.

Using the affinity binding reagents described above, the present method may be carried out, in one preferred embodiment, as follows. The high affinity binding reagents, each containing a unique affinity binding domain and being present in a mixture of anywhere from 1 to 500,000 (each with equilibrium constants of less than 500 nM), are combined with a sample (for example, a biological sample), under conditions which allow each affinity binding domain to reproducibly recognize a binding partner(s). Following complex formation, the complex is captured. This can be accomplished through any standard procedure, for example, by biotinylation of the biological sample, followed by capture of biotinylated complexes using immobilized streptavidin (for example, streptavidin immobilized on magnetic beads or a column). Alternatively, the initial protein sample may be preabsorbed onto a membrane and the binding domains mixed with the membrane. Complexes remain bound, while unbound binding reagents are washed away.

Following capture of bound complexes, binding domains that have bound their target(s) in the biological sample are detected simply by performing a PCR reaction using primers which hybridize to the nucleic acid portion of the fusion molecule. Preferably, the PCR reaction is carried out using standard quantitative methods (for example, using Taq Man by Perkin-Elmer).

If multiple complexes are isolated, the isolated pool is then deconvoluted and individual members identified. The identification step may be accomplished through direct sequencing. Alternatively, in a preferred feature of the invention, the isolated pool is deconvoluted and bound analytes

identified using DNA chip array detection. In one preferred method, the PCR reaction is stopped following predefined cycles, and aliquots extracted. In this way, DNA array detection is performed on each aliquot, allowing for quantitative analysis of amounts of each species present in the pool. Again, a critical feature of the PCR step is that the unique identifiable tag is amplified, and that each amplified segment is the result of using identical primers that generate a DNA product of identical (or essentially identical) size.

In addition, the present invention includes kits for carrying out any of the methods described herein. Typically, these kits include at least three important components: (a) a nucleic acid-protein fusion having a protein portion that specifically binds to a desired compound and a nucleic acid portion that both encodes the protein portion and includes a unique identification tag; (b) a PCR primer pair, in which the first primer is designed to hybridize to the nucleic acid portion of the fusion 5' to the unique identification tag and the second primer is designed to hybridize to the nucleic acid portion of the fusion 3' to the unique identification tag and in which hybridization of the primers to the fusion permits amplification of the unique identification tag; and (c) a solid support that includes a nucleic acid that can hybridize to the unique identification tag

In preferred embodiments, the kits of the invention may include any of the preferred components discussed above with respect to the preferred methods of the invention.

Use

The reagents of the present invention have a multitude of uses in the diagnostic industry and may be substituted for monoclonal antibodies in any use or method for which such antibodies are employed. In addition, the

molecules described herein may be used for deconvoluting species in complex biological samples. In addition, the present approach offers a fast, simple way for detecting species using PCR, a detection technique that is virtually unparalleled in its sensitivity. Coupled with the unique advances for PCR in quantitation of individual species and in DNA array platforms for deconvoluting and detecting low levels of species, the present invention represents the state of the art in the diagnostic area.

Other Embodiments

Other embodiments are within the claims.

All publications mentioned in this specification are hereby incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

What is claimed is:

Claims

1. A method for detecting multiple compounds in a sample, said method comprising:

(a) contacting said sample with a mixture of binding reagents, said binding reagents being nucleic acid-protein fusions, each having (i) a protein portion which is known to specifically bind to one of said compounds and (ii) a nucleic acid portion which encodes said protein portion and which includes a unique identification tag;

(b) allowing said protein portions of said binding reagents and said compounds to form complexes;

(c) capturing said binding reagent-compound complexes;

(d) amplifying said nucleic acid portions of said complexed binding reagents; and

(e) detecting said unique identification tag of each of said amplified nucleic acids, thereby detecting the corresponding compounds in said sample.

2. The method of claim 1, wherein said sample is a biological sample.

3. The method of claim 1, wherein said nucleic acid-protein fusion is an RNA-protein fusion.

4. The method of claim 1, wherein said nucleic acid-protein fusion is covalently bound.

5. The method of claim 4, wherein said nucleic acid-protein fusion is covalently bound through a peptide acceptor.

6. The method of claim 5, wherein said peptide acceptor is puromycin.

7. The method of claim 1, wherein said binding reagents do not bind said compounds through compound-specific antibody domains.

8. The method of claim 1, wherein each of said binding reagents comprises a scaffold domain.

9. The method of claim 8, wherein each of said binding reagents comprises a fibronectin scaffold domain.

10. The method of claim 9, wherein said fibronectin scaffold domain is the 10^{th} domain of fibronectin type III.

11. The method of claim 8, wherein each of said binding reagents comprises an antibody scaffold domain.

12. The method of claim 1, wherein said binding reagents bind said compounds with equilibrium constants of less than about 500 nM.

13. The method of claim 1, wherein said unique identification tags are detected using a solid support to which are immobilized nucleic acids specific for said unique identification tags and said detection is accomplished by hybridization of said unique identification tags to said immobilized nucleic acids.

14. The method of claim 1, wherein said amplifying step (d) is carried out using quantitative PCR.

15. The method of claim 1, wherein said compounds are proteins.

16. The method of claim 1, wherein said mixture of binding reagents comprises at least 5 different nucleic acid-protein fusions, each specifically binding to a different compound.

17. The method of claim 16, wherein said mixture of binding reagents comprises at least 100 different nucleic acid-protein fusions, each specifically binding to a different compound.

18. The method of claim 17, wherein said mixture of binding reagents comprises at least 40,000 different nucleic acid-protein fusions, each specifically binding to a different compound.

19. The method of claim 18, wherein said mixture of binding reagents comprises at least 500,000 different nucleic acid-protein fusions, each specifically binding to a different compound.

20. A method for detecting a compound in a sample, said method comprising:

(a) contacting said sample with a binding reagent, said binding reagent being a nucleic acid-protein fusion having (i) a protein portion which is known to specifically bind to said compound and (ii) a nucleic acid portion which encodes said protein portion and which includes a unique identification tag;

(b) allowing said protein portion of said binding reagent and said compound to form a complex;

(c) capturing said binding reagent-compound complex;

(d) amplifying said nucleic acid portion of said complexed binding reagent; and

(e) detecting said unique identification tag of said amplified nucleic acid, thereby detecting the corresponding compound in said sample.

21. A kit for carrying out compound detection, said kit comprising:

(a) a nucleic acid-protein fusion, wherein said protein portion of said fusion specifically binds said compound and said nucleic acid portion of said fusion encodes said protein portion and includes a unique identification tag;

(b) a PCR primer pair, wherein the first of said primers hybridizes to said nucleic acid portion of said fusion 5' to said unique identification tag and the second of said primers hybridizes to said nucleic acid portion of said fusion 3' to said unique identification tag and hybridization of said primers to said nucleic acid fusion permits amplification of said unique identification tag; and

(c) a solid support comprising a nucleic acid which can hybridize to said unique identification tag.

22. The kit of claim 21, wherein said kit further comprises Taq polymerase.

23. The kit of claim 21, wherein said nucleic acid-protein fusion is an RNA-protein fusion.

24. The kit of claim 21, wherein said nucleic acid-protein fusion is covalently bound.

25. The kit of claim 24, wherein said nucleic acid-protein fusion is covalently bound through a peptide acceptor.

26. The kit of claim 25, wherein said peptide acceptor is puromycin.

27. The kit of claim 21, wherein said nucleic acid-protein fusion does not bind said compound through a compound-specific antibody domain.

28. The kit of claim 21, wherein said nucleic acid-protein fusion comprises a scaffold domain.

29. The kit of claim 28, wherein nucleic acid-protein fusion comprises a fibronectin scaffold domain.

30. The kit of claim 29, wherein said fibronectin scaffold domain is the 10^{th} domain of fibronectin type III.

31. The kit of claim 28, wherein said nucleic acid-protein fusion comprises an antibody scaffold domain.

32. The kit of claim 21, wherein said nucleic acid-protein fusion binds said compound with an equilibrium constant of less than about 500 nM.

33. The kit of claim 21, wherein said solid support is a chip.

34. The kit of claim 21, wherein said solid support comprises an ordered array of single-stranded nucleic acids on its surface, each of said single-stranded nucleic acids being capable of hybridizing to a different said unique identification tag.

35. The kit of claim 21, wherein said compound is a protein.

36. The kit of claim 21, wherein said kit comprises at least 5 different nucleic acid-protein fusions, each specifically binding to a different compound.

37. The kit of claim 36, wherein said kit comprises at least 100 different nucleic acid-protein fusions, each specifically binding to a different compound.

38. The kit of claim 37, wherein said kit comprises at least 40,000 different nucleic acid-protein fusions, each specifically binding to a different compound.

39. The kit of claim 38, wherein said kit comprises at least 500,000 different nucleic acid-protein fusions, each specifically binding to a different compound.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/00291

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68; C12P 19/34; C07H 21/00, 21/02, 21/04
US CL : 435/6, 91.1, 91.21; 536/23.1, 23.4

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6, 91.1, 91.2, 91.21; 536/23.1, 23.4

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 98/31700 A1 (THE GENERAL HOSPITAL CORPORATION) 23 July 1998 (23.07.98), see entire document, especially pages 2-7.	1-39
Y	ROBERTS et al. RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc. Natl. Acad. Sci., USA. November 1997, Vol. 94, pages 12297-12302, see entire document.	1-39
Y	WO 99/51773 A1 (PHYLOS, INC.) 14 October 1999 (14.10.99), see entire document.	1-39

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
2/23/01

Date of mailing of the international search report
04 MAY 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Alexander H. Spiegler
Telephone No. (703) 305-3257

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/00291

Continuation of B. FIELDS SEARCHED Item3:

WEST, EPO, JPO, STN (Medline, Biosis, Embase, CAPlus, Biotechds)
search terms: RNA-protein fusion, covalent, amplify/amplification, peptide acceptor, puramycin, identification tag, label, solid support